Winning At Russian Roulette ELLSBERG

There is no denying that ~~the~~ counterexamples ~~thatxhave~~ that have ~~been~~ to the von Neumann-Morgenstern utility axioms/turned up so far ~~have been~~ are rather an odd lot. One spends his evenings at Russian roulette; another won't bet a dime to a dollar that the sun will rise tomorrow. Their fellows are the mountaineer who forsakes the Catskills for the Matterhorn, the slot-machine player who holds a dollar's worth of nickels and means to play till he has none.

Should a theory that purports to describe all "reasonable" behavior applly to ~~these~~ such people or not? Few proponents of the utility axioms, aside from Marschak[1] dismiss exceptions a priori as "unreasonable" or "mistaken." And they expect that even reasonable subjects will show "inconsistency" when they make decisions hastily; indeed, the literature abounds in "test-yourself quizzes likely to reveal violations in the reader"s snap choices. But the utility-theorists are confident that they could pursuade an educated person to "correct" such decisions, at least if he~~r~~ were given "a couple of years to brood over his answers."

~~The To challenge the theory squarely,~~ Such violators as the four above do challenge this ~~viewxinsofa~~ belief ~~as~~ inasmuch as their choices do not seem merely careless; reflection, or a course in game-theory, might not alter them. Yet they do not ~~seem to~~ contradict strongly the view expressed by Strotz: "In my own opinion, it would be a strange man indeed who would persist in violating these precepts once he understood clearly in what way he was violating them." They are pretty strange. Russian roulette players may not be crazy, but it does seem that a working theory of rational choice can afford to ignore them.[3]

I propose to exhibit ~~some violators~~ persistent violators of the axioms that ~~some~~ utility-theorists may find harder to dismiss as eccentric. The utility axioms define rational choice under risk, i.e., when the probabilities of possible consequences are known. For choices under uncertainty (i.e., when probabilities are unknown or undefined), some of these same theorists have ~~an~~ endorses alternative notions of rational behavior. To the extent that these principles are acceptable, people who follow them in choosing under uncertainty are "certified rational." It is among them one may look for counterexamples.

It turns out that in situations that combine both risk and uncertainty, the two notions of rationality may conflict. In particular: precisely those ultraconservative individuals who are rational inthe sense of the von Neumann-Morgenstern theory of games may violate the von Neumann-Morgenstern axioms on utility.

Such an argument, of course, cuts two ways.

One man spends his evenings at Russian Roulette; another won't bet a dime to a dollar that the sun will rise tomorrow. Both violate the von Neumann-Morgenstern utility axioms.

They have companions: the mountaineer who forsakes the Catskills for the Matterhorn, the slot-machine player who holds a dollar's worth of nickels and means to play till he has none. Among counterexamples to theories of rational choice, these four enjoy a special status. If we reserve judgement on the first, they are not obviously mad, or even careless. Their choices do not seem to be in the nature of *mistakes*;[1] reflection, or a course in game theory, might not alter them. We may say to the slot-machine player: "Don't you know you are *inconsistent*?" This gives him pause. He asks: "With what?"

"Why, with Axiom 3:C:b and Marschak's Theorem 6!" He puts another nickel in.

Samuelson reports a similarly unhelpful attitude:

> "The most rational man I ever met, whom I shall call Ysidro, determined his own ordinal preference pattern...
>
> "When told that he did not satisfy all of the von Neumann-

---

[1] Thus, they challenge Marschak's belief that behavior conflicting with the axioms is "important to economics in the same way in which...inability to measure, to imagine, to remember, are important. We treat it in the same way that we treat the fact that people make mistakes of arithmetic and infringe upon the rules of conventional logic." (Jacob Marschak, "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica*, May, 1949, pp. 139-40.)

Morgenstern axioms, he replied that he thought it more rational to satisfy his preferences and let the axioms satisfy themselves."[2]

~~Such unrepentant rebels raise the issue:~~ Should a theory that purports to describe all "reasonable" behavior apply to these people or not? Few proponents of the utility axioms find it so easy as Marschak to dismiss all exceptions as a priori "unreasonable" or "mistaken."[3] But they generally agree, on the basis of their introspection that non-conformists are abnormal: at least, those who have had time to reflect on their choices (Samuelson is willing to give one "a couple of years to brood over his answers" before committing himself--and testing the theory[4]). ~~In this view,~~ As one writer has put it, "In my own opinion, "It would be a strange man indeed who would persist in violating these precepts once he understood clearly in what way he was violating them."[5]

To counter such a statement, we may confront the person who makes it with our four respectably-dressed violators. The trouble is, he is all too likely to comment: "Well, they look pretty strange to me." As a matter of fact, it is hard to disagree with ~~this~~ him. ~~There is no denying that these particular specimans have the appearance of eccentrics.~~ Russian roulette players may not be crazy, but it does seem that a practical theory of rational choice could afford to overlook them.[6]

---

[2] Paul A. Samuelson, "Probability and the Attempts to Measure Utility," Economic Review (Tokyo, Japan), July, 1950, pp. 169-70. It should be mentioned that Samuelson, who can be as inconsistent as the next fellow, has changed his mind about Ysidro. (~~See~~, "Probability, Utility, and the Independence Axiom," Econometrica, October, 1952, pp. 670-78).

[3] Like rules of logic and arithmetic, says Marschak, the utility axioms serve to describe the behavior of men "who, it is believed, cannot be 'all fools all the time.'" Marschak, op. cit., p. 111.

[4] Samuelson, "Probability, Utility, and the Independence Axiom," p. 278. The utility-theorists do expect inconsistencies to show up in hasty decisions, even with reasonable subjects. The literature abounds in "test-yourself" quizzes likely to reveal violations in the reader's unconsidered choices, but the utility-theorists are confident (again, on the basis of introspection) that they could persuade an ~~ed~~ educated person to "correct" such decisions.

[5] Robert H. Strotz, "Cardinal Utility," American Economic Review, May, 1953, pp. 384-397.

The problem for a critic of the utility theory is to produce a persistent violator who can not be so plausibly dismissed as eccentric. ~~To solve this it would seem necessary~~ But how to predict the intuitive reactions of a utility-theorist? ~~but~~ There is a clue. The utility axioms define rational choice under risk, i.e., when the probabilities of the possible outcomes are known. For choices under uncertainty (i.e., when probabilities are unknown or undefined), many of these same theorists have endorsed alternative notions of rational behavior. Thus, people who follow certain patterns of choice in this new region are, in some degree, "certified rational"; it is among them ~~we will~~ one may look for counterexamples.

It turns out that in situations that combine both risk and uncertainty, the two notions of rationality may conflict. In particular: precisely those ultra-conservative individuals who are rational in the sense of the von Neumann-Morgenstern theory of games may violate conditions for the existence of a von Neumann-Morgenstern utility index.

---

[6] A rationale for the sport's appeal is suggested below, along with some hints for improving one's game. But even this boost seems unlikely to make its fans statistically significant, given normal attrition.

The axioms in question[7] imply restrictions on "rational"[8] preferences

---

[7] John von Neumann and Oskar Morgenstern, The Theory of Games and Economic Behavior, Princeton, 1944, pp. . Marschak, op. cit., presents an alternate set. For exposition of the axioms, see: Armen A. Alchian, "The Meaning of Utility Measurement," American Economic Review, March, 1953, pp. 26-50; Robert H. Strotz, op. cit.; Daniel Ellsberg, "Classic and Current Notions of Measurable Utility," Economic Journal, September, 1954. All of these writers discuss the counterexamples mentioned at the beginning of this paper.

---

[8] This interpretation of the von Neumann-Morgenstern and the later axiom-sets has been almost universal, and will be assumed here. It is worthy of note that von Neumann and Morgenstern themselves deliberately refrained from associating the notion of "rationality" with their axioms on risk-choice, to which they assigned no normative significance.

---

among risky prospects, i.e., sets of alternative consequences with known probabilities of occurrence. Prospects are known metaphorically as "lottery tickets" in which the alternative outcomes are "prizes." If A and B are prizes, the lottery ticket offering either A with probability p or B with probability (1-p) may be symbolized: (A,B; p). The axioms imply that the individual has a complete preference ordering over all prizes and lottery tickets, and they specify certain restrictions on his preferences among lottery tickets.

For example, the von Neumann-Morgenstern axiom 3:B[9] requires

---

[9] Von Neumann and Morgenstern, op. cit., p. .

---

that if A is preferred to B, then A will be preferred to any prospect (A,B;p) where p is less than 1. This property of the preference-system is known as "monotonicity."

In later formulations of the axiom-set, a key position has been ascribed to what Samuelson calls the Special Independence Assumption.[10]

---

[10] Samuelson, "Probability, Utility, and the Independence Axiom," p. 672.

~~In one form~~ This specifies, in part, that an individual indifferent between A and B must be indifferent between (A,C;p) and (B,C; p) for any C and p (the same in both tickets).

If these and other axioms are fulfilled, it follows that a set of numbers, unique up to a linear transformation, can be assigned to prizes so that the subject's actual preferences among lottery tickets will correspond to their mathematical expectations, computed on the basis of these "utility" numbers.

In formal versions of the theory the "prizes," "outcomes" or "consequences" are primitive notions, not defined explicitly. They have been variously interpreted as amounts of money or other goods, as "market-baskets" of goods or more general vectors, and as other lottery tickets. No ~~written~~ discussion has ruled out more general interpretations.[11]

---

[11] In fact, the interpretation below--that the "outcomes" ~~are~~ may be from
vectors of alternatives without known probabilities--follows directly/
various suggested applications of the theory. ~~in various discussions. This
point is developed below.~~

I shall consider, first, prospects which offer as outcomes strategies in a game. Such prizes can be regarded neither as "sure outcomes" nor as other prospects; it is axiomatic in game-theory that the set of alternative possible outcomes associated with a strategy have no assigned probabilities. A lottery ticket with strategies for prizes corresponds to the game-theoretical notion of a "mixed strategy."

Although the approach does not arise in game-theory, we can easily imagine offering a player the opportunity of playing one of a restricted set of strategies in a particular game. To understand what this might mean, it is useful to think of a man engaged in 100 plays of a game simultaneously, like a checker champion who takes on a roomful of opponents at once. Of all the possible strategies and combinations of strategies he might use, in this "super-game" we may allow him just two choices: strategy A or strategy B, the same strategy to be used in all 100 games. This would correspond to a forced choice between two "pure" strategies. On the other hand, we might allow him to flip a coin for each play to decide whether to use A or B in that play. This would be a lottery ticket, the "super-game" mixed strategy (A,B; $\frac{1}{2}$). If he chose it, he would end up using each strategy in roughly fifty plays, in random order. ¶ Another interpretation, closer to the spirit of the utility axioms, would refer not to a super-game but to a single, isolated play. The lottery ticket (A,B; $\frac{1}{2}$) offers a prospective player the following prizes; if a fair coin comes up heads, he must play strategy A in a particular game; if it comes up tails, strategy B. We will assume that the player has definite preferences among pure strategies and such lottery tickets.[12]

Now let us consider a man who is "rational" in the sense of the von Neumann-Morgenstern theory of games. In choosing among strategies, he "maximins"; i.e., he orders them with respect to their minimum elements, and he prefers the strategy with the highest minimum. Among prospects, he maximizes the mathematical expectation of utility.[13]

---

[12] Specifically, we assume the following: if the player is restricted to any finite set of pure or mixed strategies, there will exist a subset which are preferred to all others in the set but not to each other. A complete ordering is not assumed, though it follows from a later assumption.

[13] For the moment, we may assume that he has a von Neumann-Morgenstern utility index over "sure outcomes," derived by observing his choices among prospects which offered only sure outcomes or other prospects as prizes.

This implies that he will derive the outcomes associated with the mixed strategy (X,Y;$\frac{1}{2}$) by averaging corresponding pairs of payoffs for X and Y. In other words, he will be indifferent between the mixed strategy and a pure strategy with outcomes equal to these average payoffs.

If we offer him the strategies A,B and C shown in Figure 1 (where a and b are the "opponent's strategies"), this player will prefer A to B, B to C, and A to C. Now we ask him to consider the prospect (B,C;$\frac{1}{2}$); by assumption, he will derive its payoffs by averaging the corresponding payoffs for B and C. Although he prefers strategy A to either B or C, he will report that he prefers a lottery ticket offering either B or C to the certainty of A! Similarly, preferring A to B, he nevertheless prefers the prospect (A,B; $\frac{1}{2}$) to the certainty of A. The choices clearly violate the monotonicity condition.

Figure 1.

| | a | b |
|---|---|---|
| A | 3 | 6 |
| B | 7 | 2 |
| C | 1 | 8 |
| (A,B;$\frac{1}{2}$) | 5 | 4 |
| (B,C;$\frac{1}{2}$) | 4 | 5 |

Figure 2.

| | a | b |
|---|---|---|
| A | 3 | 6 |
| A' | 6 | 3 |
| B | 7 | 2 |
| C | 1 | 8 |
| (A,B;$\frac{1}{2}$) | 5 | 4 |
| (A',B;$\frac{1}{2}$) | 6.5 | 2.5 |
| (A',C;$\frac{1}{2}$) | 3.5 | 5.5 |

Figure 2. indicates that Samuelson's Special Independence Assumption fares no better. The player who is rational in the game-theoretical sense is indifferent between A and A'; but he positively prefers (A,B;$\frac{1}{2}$) to (A',B$\frac{1}{2}$). Moreover, although he prefers B to C, he prefers (C,A';$\frac{1}{2}$) to (B,A';$\frac{1}{2}$).

The game in Figure 3. illustrates another contradiction of the monotonicity condition. A minimaxer prefers A to B, but he prefers (A,B; 2/100) to (A,B;$\frac{1}{2}$) although the probability of the preferred alternative is higher in the latter.[14]

Figure 3.

|  | a | b |
|---|---|---|
| A | 1 | 99 |
| B | 2 | 0 |
| (A,B;$\frac{1}{2}$) | 1.5 | 49.5 |
| (A,B;2/100) | 1.98 | 1.98 |

For this man it is possible to define an ordinal utility function over strategies such that he can be said to be maximizing utility. For one such index, simply take the minimum payoff for each strategy (he has a complete ordering over strategy-prizes). Any positive monotone transformation of this index will do as well. But among these ordinal utility indices there will not exist a "von Neumann-Morgenstern utility index." It is not possible to assign numbers to strategies in such a way that in his choices among lottery tickets with strategies for prizes he maximizes the mathematical expectation of these utility numbers.[15] He violates the axioms.

---

[14] This result has some plausibility in games against informed opponents. But when the minimax principle is followed in "games against Nature" (discussed below) such preferences are somewhat paradoxical.

[15] It is not necessary to these results that he have a strict von Neumann-Morgenstern utility index over "sure outcomes" either. In terms of the example in Figure 1, no matter how the payoffs are defined, just so long as the player prefers a 50-50 chance of 7 or 1, and of 2 or 8, to the certainty of 3, he will prefer (A,B;$\frac{1}{2}$) to A and our conclusions stand.

Prospects of the sort we have been considering have more than formal existence. They can be observed wherever the concept of a mixed game-strategy is appropriate. With certain modifications, our conclusions would extend to prospects offering games as prizes. This too has its real (though not game-theoretical) counterparts. In poker, players may flip a coin to decide whether to play stud or draw, or whether to allow sandbagging; and in the business world, similar lotteries may be introduced in the course of "bargaining over the rules of the game." Still, a proponent of the utility axioms might maintain that in practice, such situations could be objectively distinguished from prospects with more conventional prizes. He might be willing to exclude prospects of strategies ~~of~~ from the application of the axioms.[16]

This limitation is much more severe than our previous discussion might suggest, since it would apply not only to strategies against rational opponents but to strategies and mixed strategies in "games against Nature." That ~~This consideration~~ enlarges enormously the set of situations combining risk and uncertainty to which the axioms would not~~a~~ apply. At the same time, it introduces new types of "reasonable" behavior which would violate the axioms if the latter were applied. In fact, nearly every criterion of choice that has been proposed as "reasonable" in "games against Nature"[17] would conflict with the

---

[16] He might add that this was his intention from the beginning: though there is no hint of this in any published discussion of the axioms.

[17] See John Milnor, "Games Against Nature," in Decision Processes, ed. Thrall, Coombs and Davis, New York, 1954, pp. 49-60.

~~with the~~ utility axioms in situations where both would apply.

In games against Nature, the consequences of a particular action are determined by "events" whose probabilities are unknown. In the situation represented by Figure 4 (where *a* and *b* are the "events"), let us suppose that the decision-maker follows the Savage principle of "minimaxing regret."[18] He bases his decision on the "regret" matrix in Figure 4b, derived by replacing each element in the payoff matrix (Figure 4a) by its numerical difference from the maximum element in its column. He chooses the action with the least maximum "regret." Like our previous examples (and, incidentally, like the mountain-climber and the player of Russian roulette) he prefers A to B but prefers the lottery ticket $(A,B;\frac{1}{2})$ to either.

*Figure 4a*

Payoff matrix

| | *a* | *b* |
|---|---|---|
| A | -8 | 0 |
| B | 0 | -10 |

*Figure 4b*

Regret matrix

| | *a* | *b* |
|---|---|---|
| A | 8 | 0 |
| B | 0 | 10 |
| $(A,B;\frac{1}{2})$ | 4 | 5 |

As a matter of fact, use of this criterion to rank each pair of alternatives in turn will lead to intransitive preferences,[19] as shown in Figure 5. Thus, for a person using this principle, not even an ordinal utility index over strategies will exist, much less a von Neumann-Morgenstern index.[20]

---

[18] Milnor, *op. cit.*, p. 50.

Chernoff

[19] This fact has gone unremarked in the literature. Milnor's puzzling assertion (*op. cit.*, p. 52) that minimaxing regret yields a complete ordering apparently refers to transitivity within a particular regret matrix. It seems much more natural to test for transitivity of the criterion in pairwise comparisons.

[20] Of course, this would also follow from the familiar fact that minimaxing regret violates the "independence of irrelevant alternatives." Milnor, *op. cit.* p. 52.

Figure 5a. Payoff matrix

| | a | b | c |
|---|---|---|---|
| A | 4 | 0 | 0 |
| B | 0 | 3 | 2 |
| C | 3 | -1 | 2 |

Figure 5b. Regret matrices

| | a | b | c | |
|---|---|---|---|---|
| A | 0 | 3 | 2 | |
| B | 4 | 0 | 0 | A(p)B |
| B | 3 | 0 | 0 | |
| C | 0 | 4 | 0 | B(p)C |
| A | 0 | 0 | 2 | |
| C | 1 | 1 | 0 | C(p)A |

A third rule for choice under uncertainty is known as the Hurwicz criterion.[21] It implies that the individual will rank the actions according to the index $\alpha A + (1-\alpha)a$, where A is the maximum element in the row, a is the minimum element and $\alpha$ is a constant parameter $(0 \leq \alpha \leq 1)$ indicating the "degree of optimism." In Figure 4, for any $\alpha$ the man will prefer A to B but, once again, will prefer a 50-50 combination of A or B to either.

The only criterion that has been proposed for such situations that will not lead to violations of the utility axioms is what Milnor calls the "Laplace" criterion.[22] This says, where probabilities are unknown, assume they are all equal; assign equal probabilities to the alternatives outcomes of each action and maximize the weighted sum of outcomes. Familiar ~~as~~ in connection with the "principle of insufficient reason," this criterion has well-known shortcomings,[23] ~~though possibly this new aspect will count in its favor for some.~~ Since it will never score a mixed strategy higher than all of its components, it cannot lead to the contradictions we have been considering.

---

[21] ibid., p. 50.

[22] ibid., p. 49.

[23] ibid., p. 56.

The availability of the Laplace criterion hardly suffices to rule out the other, non-probabilistic criteria, either as descriptive hypotheses or as normative guides to "reasonable" behavior. Neither do Savage's recent axioms on rational choice,[24] which lead to assigning "personal probabilities" to events. A reader of Savage's book might well get a contrary impression, if he overlooked Savage's own remark: "There are practical circumstances in which one might well be willing to accept the ((minimax regret)) rule--even one who, like myself, holds a personalistic view of probability."[25] Among other things, he mentions that "an aura of vagueness is attached to many judgments of personal probability." It is just in such situations that the criteria we have discussed seem acceptable.

If it is ever "reasonable" to minimax loss, or minimax regret, or to use a Hurwicz criterion, then a reasonable man may make choices among prospects that violate the von Neumann-Morgenstern axioms. To be sure, this holds "only" for prospects that involve uncertainty as well as risk. *But what prospect does not*? It may be that for certain minor, short-run decisions the uncertainties may safely be disregarded, or reduced to probability distributions. But consider Marschak's characterization of the intended scope of the utility axioms:

> "If the space *X* of "outcomes" and the space *S* of strategies are defined, this permits us to take care of all human decisions, transcending conventional economics and including the private man's choice

---

[24] Leonard J. Savage, *The Foundations of Statistics*, Wiley, New York, 1954.

[25] *ibid.*, p. 168; he goes on to describe such circumstances.

of profession or wife, the legislator's choice of election tactics or national policies or military and administrative decisions."[26]

---

[26] Jacob Marschak, "Why 'Should' Statisticians and Businessmen Maximize Moral Expectation?" Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1951, p. 499.

---

The decisions he mentions seem precisely to fit our framework. The choice of a wife, ~~for example~~ indeed: the oldest game against Nature. Who will deny that his pre-nuptial ~~judgments of probability~~ expectations possessed, to put it mildly, an "aura of vagueness"? Election tactics, military decisions: these are strategies in a game. Allow sufficient "vagueness" of likelihoods to make minimax principles acceptable, and in choosing among prospects with campaigns or brides for outcomes, the legislator or prospective husband will slip out from Marschak's theory.

In fact, who will remain? If the axioms are still to apply to lotteries with only "sure consequences" (not strategies) for prizes, it no longer seems an easy problem even to decide when in reality these conditions ~~are approximated~~ do hold. As Savage puts it:

> "What are often thought of as consequences (that is, sure experiences of the deciding person)...are in reality highly uncertain. Indeed, in the final analysis, a consequence is an idealization that can perhaps never be well approximated."[27]

---

[27] Savage, op. cit., 84; see pp. 83-84 for an excellent discussion of this point.

---

~~Our discussion~~ This is not to imply that our discussion so far serves to "rationalize" the particular counterexamples introduced at the beginning of this paper; their behavior may depend on considerations far different from the ones we have discussed. Still, it may be worthwhile to take a second look at them. The mountaineer and the Russian roulette player, for example: the alternative outcomes of their actions have always been described as "Life" or "Death." yet Even the crudest simplification must ~~admit~~ allow the possibilities that Death may be associated either with Heaven or with Hell, and it is plausible that our subject may not assign probabilities to these events. To simplify the game, let us also assume that he has strong time-preference, ignoring consequences beyond the next few years (it seems likely that Russian roulette players do not tend to live in the future). If we arbitrarily choose the origin of his utility function so that the value of living a few more years is zero, the payoff function for Russian roulette may appear as in Figure 6 (where the strategy Death represents six bullets in the ~~chamber~~ cylinder, and Life, none).

Payoff matrix

Figure 6.

$$\begin{array}{lcc} & \underline{\text{Heaven}} & \underline{\text{Hell}} \\ \text{Life} & 0 & 0 \\ \text{Death} & 1000 & -5000 \end{array}$$

An appropriate criterion here would be to minimax regret. After all, the possibility of foregoing Heaven in the immediate future is a mark against Life; and the prospect of moving to Hell ~~years~~ sooner than necessary is a deterrent to Death. The player may, then, consider the regret matrix in Figure 7.

Regret matrix

Figure 7.

$$\begin{array}{lcc} & \underline{\text{Heaven}} & \underline{\text{Hell}} \\ \text{Life} & 1000 & 0 \\ \text{Death} & 0 & 5000 \\ \text{(Life, Death;5/6)} & 833 & 833 \end{array}$$

He will prefer Life to Death. But his optimum strategy is (Life, Death; 5/6). The best way for him to play is to put one bullet in the chamber and spin it.

This is, in fact, the traditional strategy. That testifies to the realism of our assumptions. But it does not follow that this choice is really best-suited to the tastes of every player. But Our rsults are not cut-and-dried. To play optimally, each person must determine for himself his own payoff function and the criterion that seems most natural to him. One player, for example, may minimax regret but place a higher value on Heaven: say, 2500. The rational course for him is to put two bullets in the chamber.

On the other hand, if after all he follows the Laplace criterion, at these payoffs he might as well live.